

**Universidade de São Paulo**  
Instituto de Ciências Matemáticas e de Computação  
SME0808 - Séries Temporais e Aprendizado Dinâmico

**Modelagem e Previsão de Séries Temporais com Tendência e  
Sazonalidade: Uma Comparação entre os Modelos de  
Holt-Winters e SARIMA**

**Lucas Alonso de Almeida - 14785493**  
**Daniel Alonso de Almeida - 14614626**  
**Igor Vieira Silva - 14675819**  
**Mie Toshio Omiya - 14572613**

São Carlos, SP  
2025

# Sumário

<b>1</b>	<b>Introdução</b>	<b>2</b>
<b>2</b>	<b>Fundamentação Teórica</b>	<b>2</b>
2.1	Modelos de Suavização Exponencial: Holt-Winters . . . . .	2
2.1.1	Holt-Winters Aditivo . . . . .	2
2.2	Modelo SARIMA . . . . .	3
<b>3</b>	<b>Metodologia</b>	<b>3</b>
3.1	Tratamento dos Dados . . . . .	3
3.2	Verificação de Estacionariedade . . . . .	4
3.3	Ajuste e Seleção de Modelos . . . . .	4
3.4	Divisão da Série e Avaliação de Desempenho . . . . .	4
<b>4</b>	<b>Resultados e Discussão</b>	<b>4</b>
4.1	Tratamento Inicial e Visualização da Série . . . . .	5
4.2	Verificação de Estacionariedade . . . . .	7
4.2.1	SARIMA . . . . .	13
4.2.2	Holt-Winters Aditivo . . . . .	17
4.3	Divisão da Série e Avaliação de Desempenho . . . . .	18
<b>5</b>	<b>Conclusão</b>	<b>21</b>

# 1 Introdução

O monitoramento e a previsão do consumo residencial de energia elétrica são fundamentais para o planejamento energético e econômico do Brasil, permitindo decisões mais informadas sobre investimentos, operação de usinas e políticas tarifárias.

Séries temporais de consumo apresentam tendência de longo prazo e sazonalidade anual, com picos no verão e inverno devido ao uso de climatização. A correta identificação desses padrões é essencial para previsões confiáveis.

Este trabalho utiliza dados mensais de consumo residencial no Brasil, na unidade de gigawatt-hora (GWh), obtidos da Empresa de Pesquisa Energética (EPE), cobrindo o período de janeiro de 2004 a julho de 2025 (259 observações). O objetivo é ajustar, diagnosticar e comparar o desempenho de dois métodos clássicos de previsão: Holt-Winters (aditivo e multiplicativo) e SARIMA, avaliando suas previsões por métricas de erro como RMSE, MAE e MAPE.

## 2 Fundamentação Teórica

Nesta seção, apresentamos os conceitos matemáticos e as formulações dos modelos de previsão utilizados neste estudo.

### 2.1 Modelos de Suavização Exponencial: Holt-Winters

O método de Holt-Winters é uma extensão da suavização exponencial simples projetada para capturar simultaneamente a tendência e a sazonalidade de uma série temporal. O método decompõe a série em três componentes: nível ( $L_t$ ), tendência ( $T_t$ ) e sazonalidade ( $S_t$ ).

#### 2.1.1 Holt-Winters Aditivo

O modelo aditivo é preferível quando a magnitude da sazonalidade não varia com o nível da série. As equações de suavização são dadas por:

$$L_t = \alpha(y_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (1)$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (2)$$

$$S_t = \gamma(y_t - L_t) + (1 - \gamma)S_{t-s} \quad (3)$$

onde  $\alpha$ ,  $\beta$  e  $\gamma$  são parâmetros de suavização entre 0 e 1, e  $s$  é o período sazonal. A previsão

$h$  passos à frente é dada por:

$$\hat{y}_{t+h|t} = L_t + hT_t + S_{t-s+h} \quad (4)$$

## 2.2 Modelo SARIMA

Os modelos SARIMA (Seasonal AutoRegressive Integrated Moving Average), denotados por  $\text{ARIMA}(p, d, q)(P, D, Q)_s$ , incorporam termos sazonais multiplicativos aos modelos ARIMA tradicionais. A estrutura geral é expressa como:

$$\Phi_P(B^s)\phi(B)(1-B)^d(1-B^s)^D y_t = \Theta_Q(B^s)\theta(B)\epsilon_t \quad (5)$$

onde:

- $y_t$  é a observação no tempo  $t$ ;
- $\epsilon_t$  é o termo de erro (ruído branco);
- $B$  é o operador de defasagem (backshift), tal que  $B^k y_t = y_{t-k}$ ;
- $d$  e  $D$  são as ordens de diferenciação simples e sazonal, respectivamente;
- $\phi(B)$  e  $\theta(B)$  são polinômios dos componentes AR e MA não sazonais de ordens  $p$  e  $q$ ;
- $\Phi_P(B^s)$  e  $\Theta_Q(B^s)$  são polinômios dos componentes AR e MA sazonais de ordens  $P$  e  $Q$ .

A identificação das ordens do modelo é tipicamente realizada através da análise das Funções de Autocorrelação (FAC) e Autocorrelação Parcial (FACP), complementada por critérios de informação como AIC (Akaike Information Criterion) e BIC (Bayesian Information Criterion).

## 3 Metodologia

A análise estatística foi conduzida utilizando a linguagem R e bibliotecas especializadas, incluindo `forecast`, `tseries` e `nortest`. O procedimento metodológico seguiu as etapas descritas abaixo:

### 3.1 Tratamento dos Dados

A série de consumo residencial foi verificada para garantir que os valores fossem tratados como numéricos e, a fim de facilitar a interpretação e o ajuste dos modelos, os valores foram escalonados dividindo-se por  $10^7$ .

### 3.2 Verificação de Estacionariedade

Para a modelagem SARIMA, a estacionariedade é um pré-requisito. Foram realizados testes de raiz unitária (ADF - Augmented Dickey-Fuller) na série escalonada. A análise indicou que uma única diferenciação ( $d = 1$ ) foi suficiente para remover a tendência estocástica da série e torná-la estacionária.

### 3.3 Ajuste e Seleção de Modelos

Para os modelos de Holt-Winters, ajustaram-se as especificações aditiva e multiplicativa utilizando a série histórica. Para o modelo SARIMA, realizou-se uma busca manual em uma grade de combinações de ordens  $(p, d = 1, q)(P, D = 0, Q)_{12}$ , estimando os parâmetros via máxima verossimilhança e avaliando os modelos através do critério BIC (calculado a partir dos resíduos dos modelos) para seleção.

### 3.4 Divisão da Série e Avaliação de Desempenho

A série temporal contempla 21 anos e 7 meses e foi dividida em dois conjuntos:

- **Treinamento:** Os primeiros 20 anos de observações, utilizado para ajuste e estimação dos coeficientes dos modelos.
- **Teste:** Composto pelos últimos 1 ano e 7 meses das observações, utilizado para validação.

O desempenho dos modelos foi avaliado com base em métricas calculadas sobre a base de teste, incluindo:

- Raiz do Erro Quadrático Médio (RMSE);
- Erro Médio Absoluto (MAE);
- Erro Percentual Absoluto Médio (MAPE).

Essa abordagem permite comparar a capacidade preditiva dos modelos e selecionar a especificação mais adequada para a série analisada.

## 4 Resultados e Discussão

Esta seção detalha o processo investigativo que percorremos, desde as primeiras impressões visuais até a modelagem formal. Nossa abordagem seguiu um fluxo natural de descoberta, onde cada observação motivou a próxima etapa de análise.

## 4.1 Tratamento Inicial e Visualização da Série

Os dados de consumo residencial de energia elétrica no Brasil (2004/01–2025/07) foram disponibilizados em GWh. Para facilitar a visualização e a análise numérica, dividimos todos os valores por  $10^7$ , mantendo a interpretação relativa do crescimento e da sazonalidade.

A série apresenta, à primeira vista:

- Tendência de crescimento clara
- Sazonalidade aparente anual, com picos típicos no verão e vales no inverno
- Possível quebra de estrutura por volta de 2014, possivelmente relacionada a fatores econômicos externos

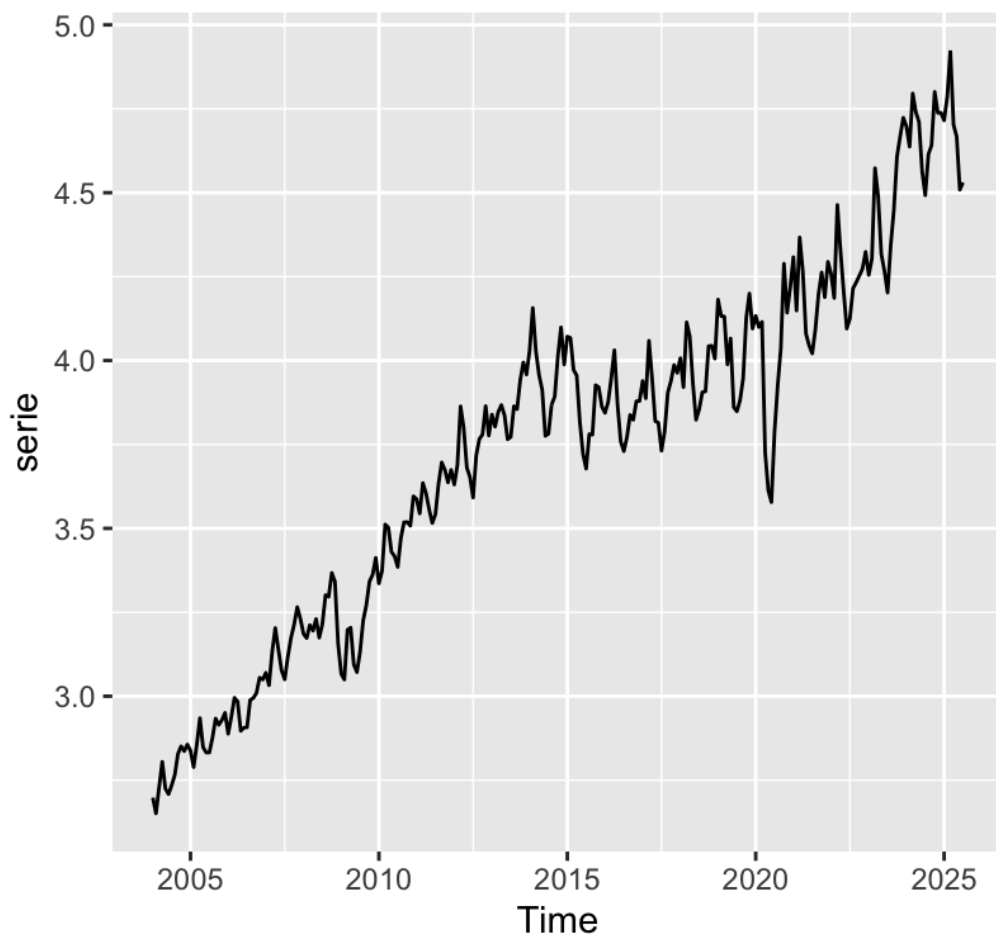


Figura 1: Série Temporal de Consumo Residencial de Energia Elétrica (escala  $10^7$  GWh). Nota-se a tendência crescente e a sazonalidade anual aparente.

A decomposição aditiva e multiplicativa da série foi realizada assumindo período sazonal de 12 meses. Surpreendentemente, ambas as decomposições apresentam componentes

muito semelhantes:

- Tendência suavizada praticamente idêntica
- Componente sazonal com padrão consistente em amplitude e fase
- Resíduos comparáveis, apesar de uma decomposição ser aditiva e a outra multiplicativa

Isso indica que a série é relativamente estável e que técnicas de suavização, como Holt-Winters, podem capturar bem os padrões gerais sem distorção significativa.

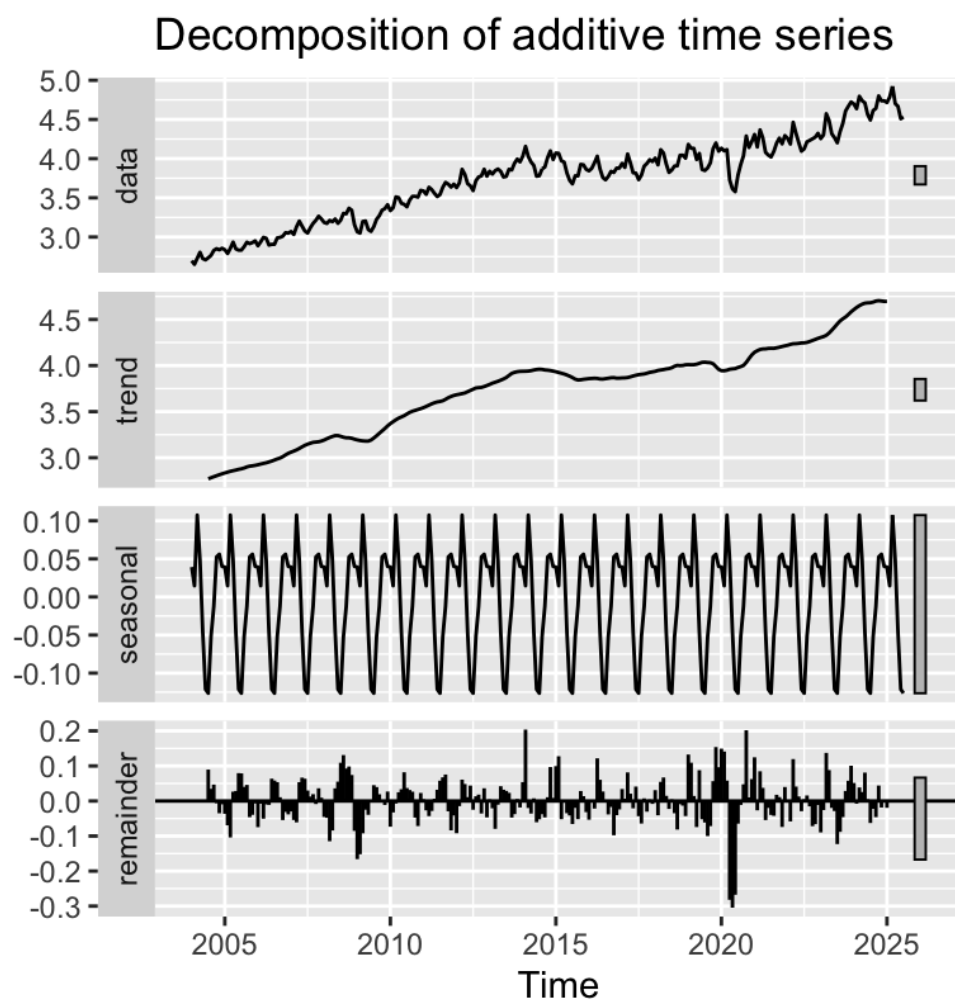


Figura 2: Decomposição aditiva da série: tendência, sazonalidade e resíduos.

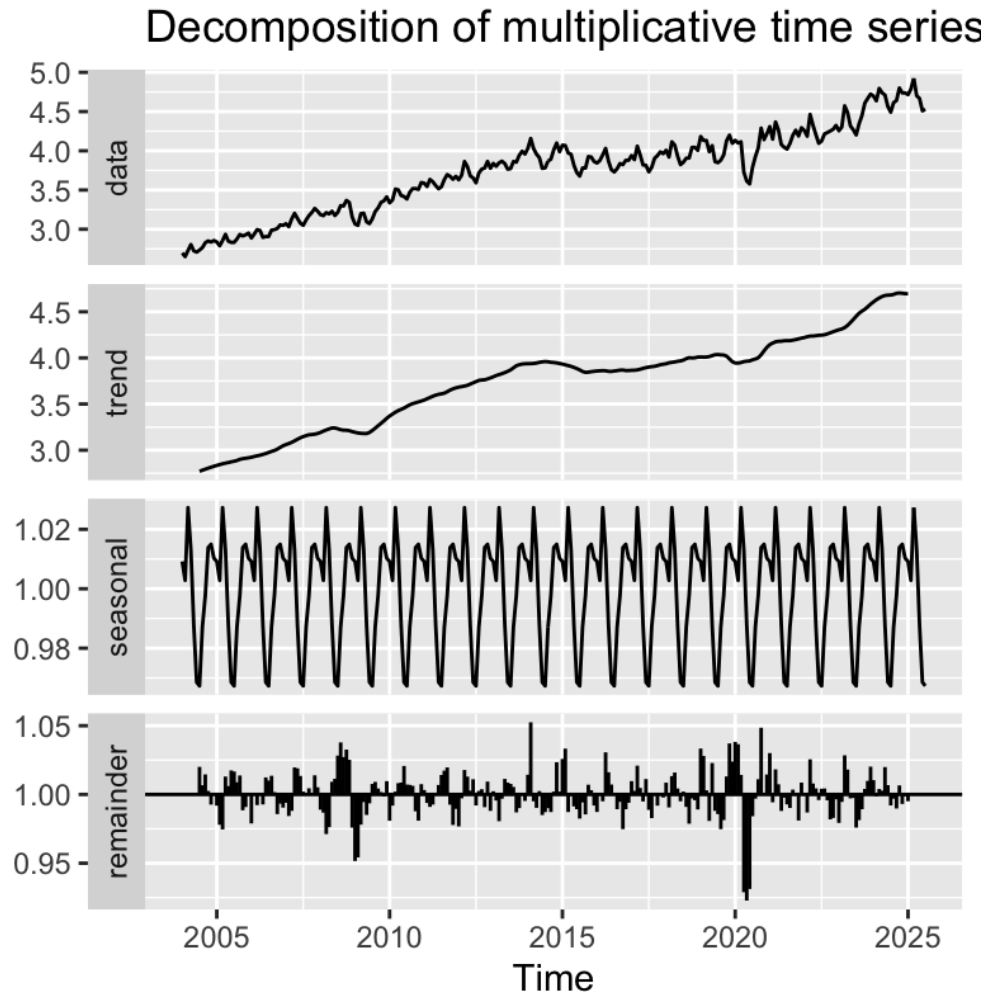


Figura 3: Decomposição multiplicativa da série: tendência, sazonalidade e resíduos. Componentes muito semelhantes à decomposição aditiva.

A análise da função de autocorrelação (ACF) da série original reforça a presença de sazonalidade anual, com picos em lags múltiplos de 12. O `seasonplot` da série evidencia novamente a tendência crescente, e dá indícios de que o ciclo anual é consistente, reforçando a escolha:

## 4.2 Verificação de Estacionariedade

Para o ajuste do modelo SARIMA, a estacionariedade da série é um requisito fundamental. Observamos uma tendência clara nos dados, característica que viola a premissa de média constante. Assim, aplicamos uma diferenciação simples ( $d = 1$ ) com o objetivo de eliminar a tendência linear e tornar a série estacionária.



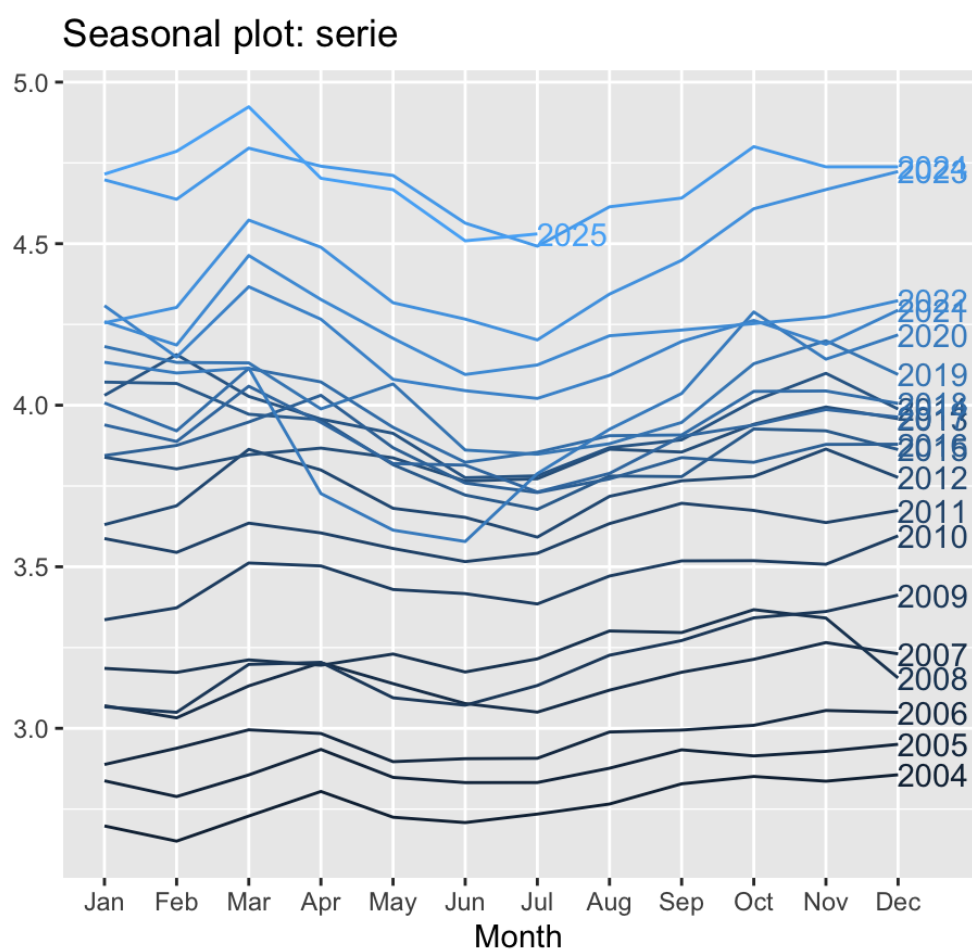


Figura 4: Seasonplot da série: destaque para a tendência ano-a-ano de crescimento e leve picos por volta de março e outubro.

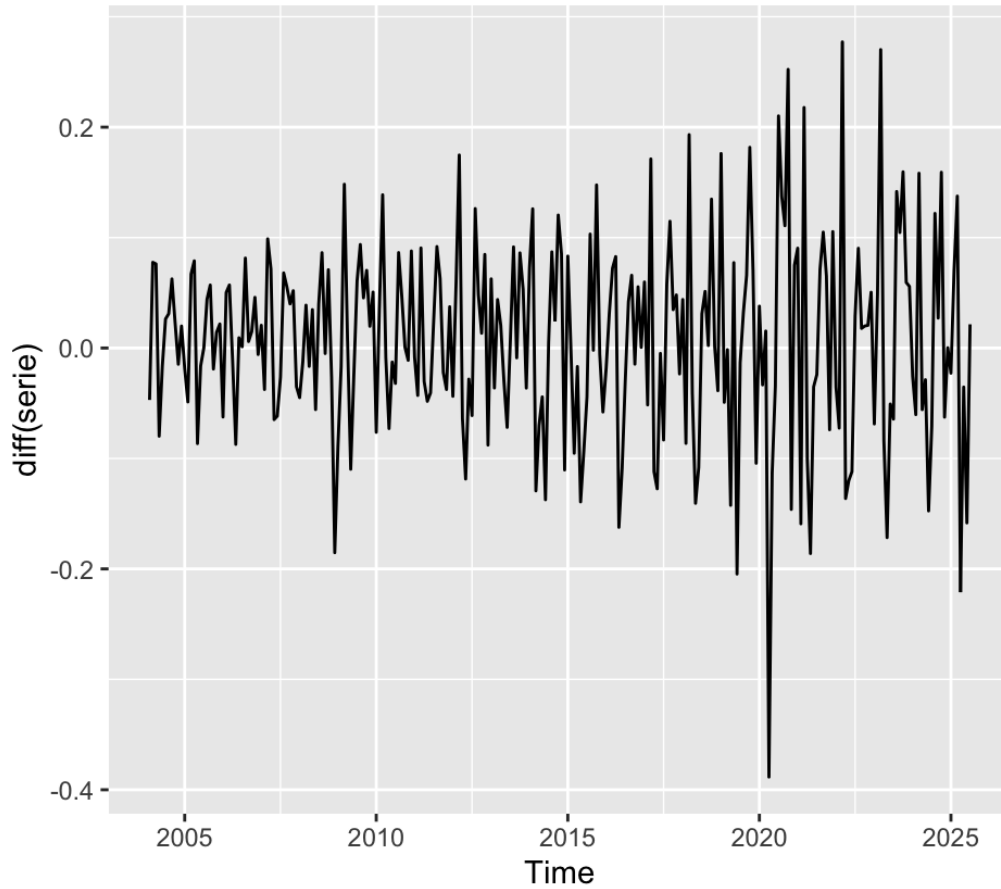


Figura 5: Série após primeira diferenciação ( $d = 1$ ). A tendência linear foi removida, e a série se aproxima de comportamento aleatório.

Ao analisar visualmente a série diferenciada, confirmamos a sazonalidade de um ano e notamos que ela se aproximava de um ruído estacionário. Para confirmar sobre o ruído, aplicamos o teste de Augmented Dickey-Fuller (ADF), que avaliou formalmente a presença de raízes unitárias.

Tabela 1: Resultados do teste ADF na série diferenciada ( $d = 1$ ).

Estatística de teste	Ordem do lag	p-valor
-8.3894	6	0.01

A hipótese alternativa do teste é a estacionariedade. O teste indicou ausência de raízes unitárias, inclusive sazonais. Isso sugere que não seria necessário aplicar uma diferenciação sazonal adicional ( $D = 1$ ). Entretanto, para maior segurança, analisamos o comportamento da função de autocorrelação (ACF) e da função de autocorrelação parcial (PACF) comparando:

- Série diferenciada uma vez ( $d = 1$ ,  $D = 0$ )

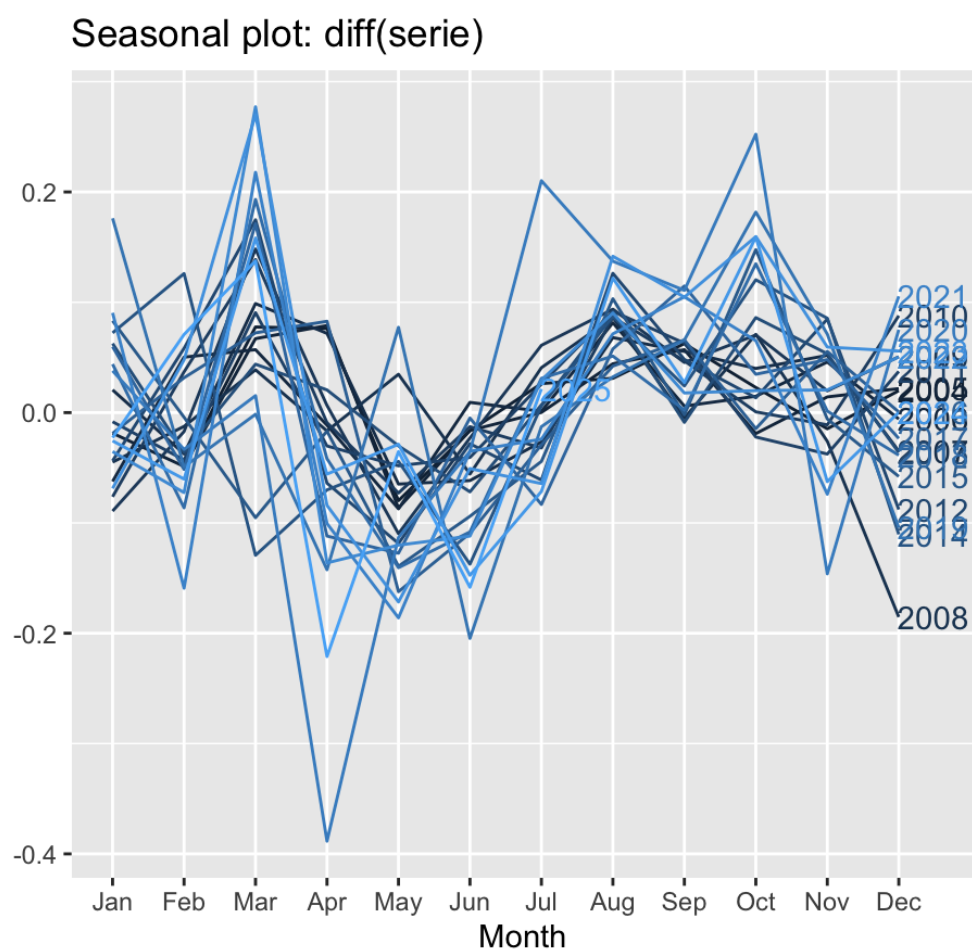


Figura 6: Seasonplot da série diferenciada: aqui os dois picos são muito mais claramente pronunciados.

- Série diferenciada duas vezes ( $d = 1, D = 1$ )

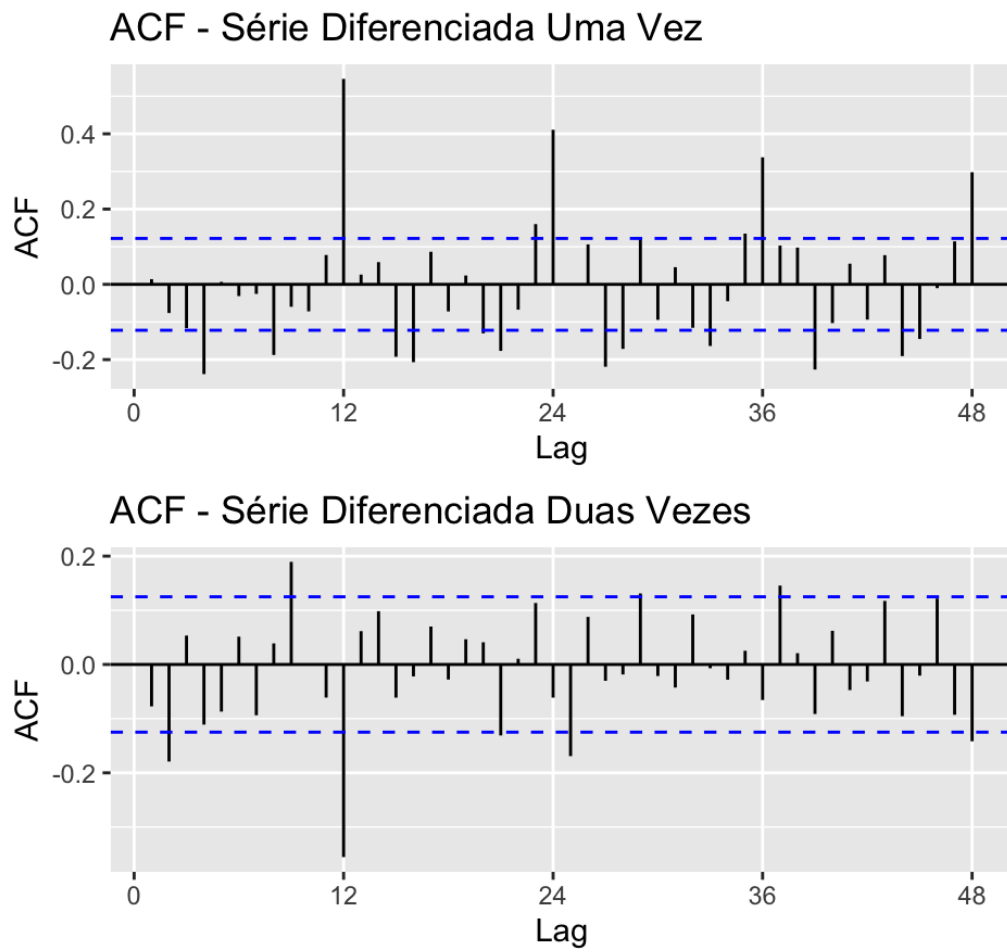


Figura 7: Comparativo da ACF entre série diferenciada uma vez e duas vezes. A série diferenciada uma vez apresenta decaimento exponencial indicativo de comportamento ARMA.

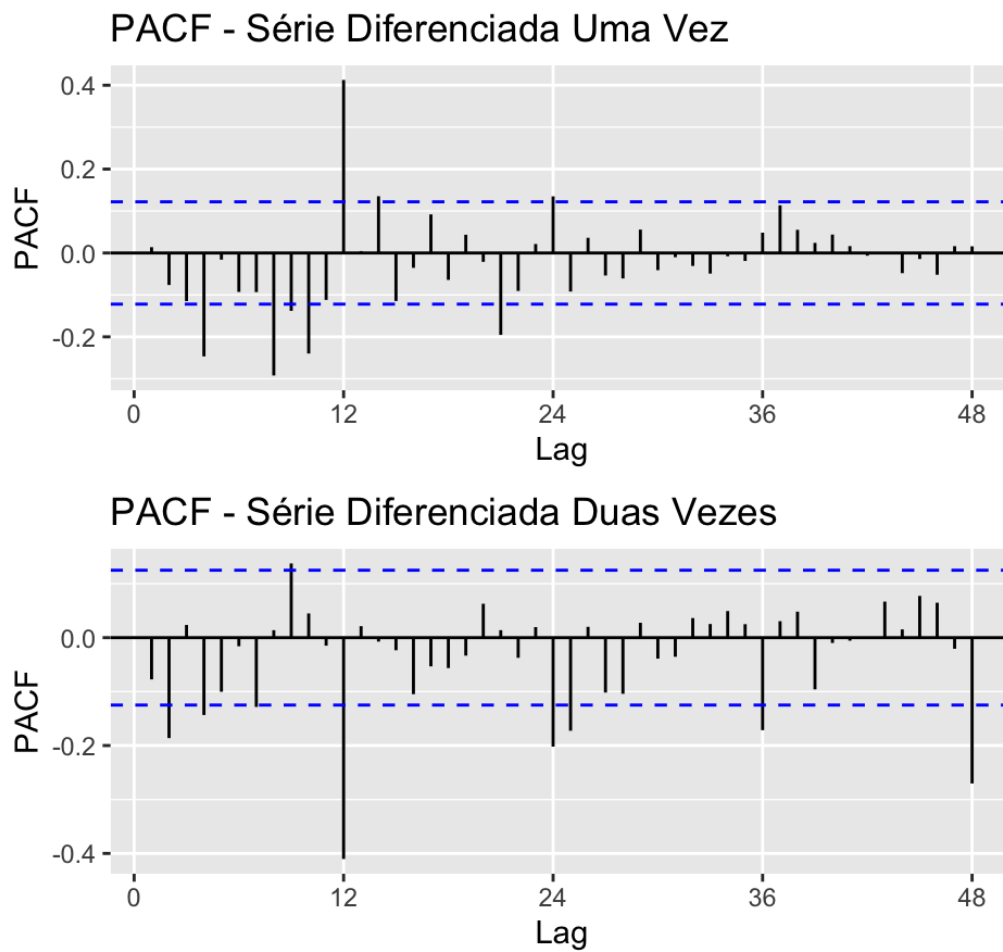


Figura 8: Comparativo da PACF entre série diferenciada uma vez e duas vezes. A série diferenciada duas vezes apresenta comportamento mais irregular, enquanto a diferenciada uma vez exibe decaimento coerente com modelo ARMA.

Para reforçar a observação visual, plotamos as séries diferenciadas lado a lado:

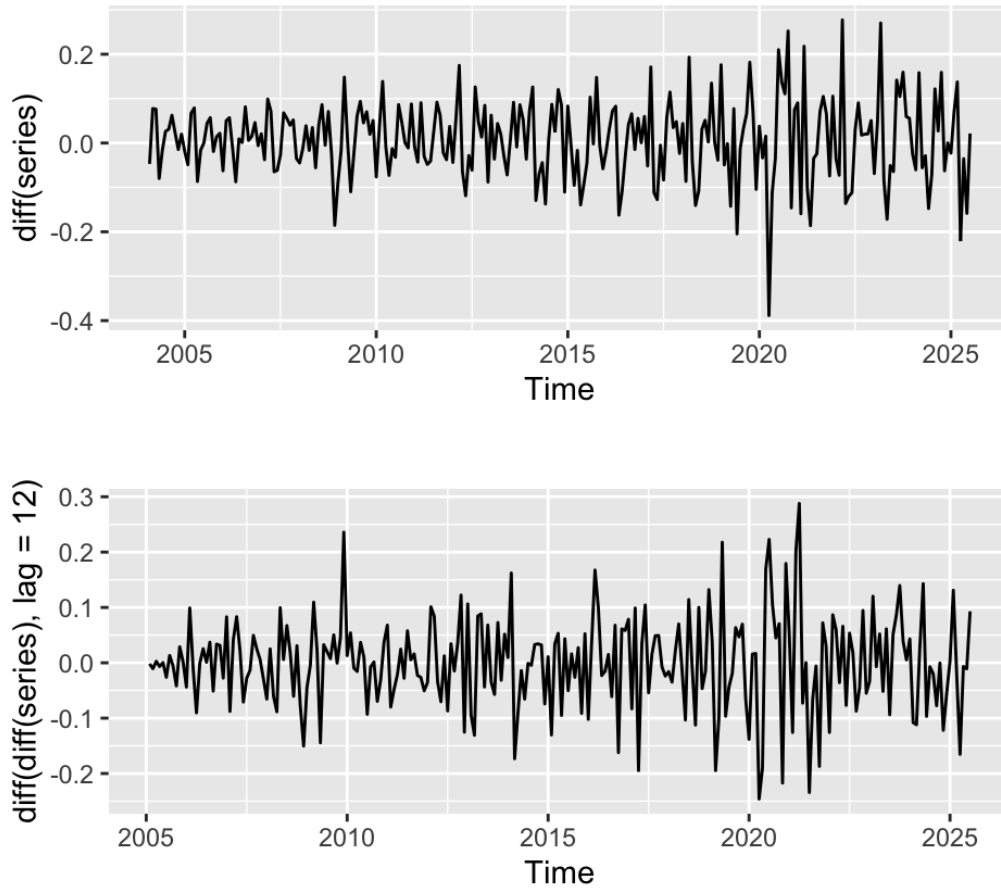


Figura 9: Comparação lado a lado da série diferenciada uma vez ( $d = 1$ ,  $D=0$ ) e diferenciada duas vezes ( $d = 1$ ,  $D=1$ ). A série diferenciada uma vez apresenta comportamento mais aleatório e adequado para modelagem SARIMA, enquanto a série diferenciada duas vezes mostra aumento de variabilidade e leve padrão senoidal.

Com base na análise visual, ACF, PACF e no teste ADF, concluímos que:

- Apenas uma diferenciação simples ( $d = 1$ ) resultou em uma série mais próxima de um ruído branco do que a série diferenciada duas vezes, que parece ter aumentado a variabilidade e introduzido um ruído senoidal.
- A série diferenciada uma vez apresenta decaimento em ACF e PACF consistente com comportamento ARMA, fornecendo base sólida para ajuste SARIMA.

#### 4.2.1 SARIMA

O ajuste do modelo SARIMA seguiu a decisão tomada na etapa de verificação de estacionariedade: apenas uma diferenciação simples ( $d = 1$ ) foi aplicada, enquanto a diferenciação sazonal não se mostrou necessária ( $D = 0$ ) segundo métodos automatizados de identificação das ordens dos componentes. A periodicidade mensal foi mantida (12 meses), e a busca exaustiva pelos melhores parâmetros foi realizada com:

- Espaço de busca:  $max.p = max.q = 3$ ,  $max.P = max.Q = 2$ ,  $max.order = 10$
- Estratégia de busca exaustiva:  $stepwise = FALSE$ ,  $approximation = FALSE$
- Critério de seleção: BIC

O algoritmo `auto.arima` testou todas as combinações possíveis dentro dessas restrições e selecionou o modelo com menor BIC, medida que penaliza modelos mais complexos. O BIC é calculado considerando a verossimilhança dos resíduos ajustados pelo número de parâmetros do modelo, favorecendo modelos parcimoniosos quando o número de observações é grande ( $n = 259$ ).

Tabela 2: Modelo ARIMA(1,1,1)(2,0,0)[12]

Parâmetro	Coefficiente	Erro Padrão
AR1	0.7435	0.0729
MA1	-0.9102	0.0432
SAR1	0.4752	0.0616
SAR2	0.1971	0.0626
$\sigma^2 = 0.004995$	log likelihood = 316.51	
AIC = -623.03	AICc = -622.79	BIC = -605.26

Como esperado, o modelo confirma que  $D = 0$ , caracterizando um ARMA sazonal ( $P = 2$ ), consistente com a análise de ACF e PACF realizada anteriormente.

Para verificar a adequação do modelo, realizamos uma análise diagnóstica.

Tabela 3: Métricas do Modelo SARIMA

Modelo	RMSE	MAE	MAPE (%)
SARIMA(1,1,1)(2,0,0)[12]	0.06999	0.05382	1.4194

**Teste de Ljung-Box** O teste indica que não há autocorrelação significativa nos resíduos, compatível com ruído branco (Tabela 4).

Tabela 4: Teste de Ljung-Box nos resíduos

Lag	Estatística	p-valor
20	27.019	0.1347

**Distribuição dos resíduos** A Figura 10 mostra o histograma dos resíduos junto com a densidade estimada e a normal teórica. Observa-se distribuição aproximadamente simétrica com leves desvios nas caudas.

## Resíduos: Histograma + Densidade Estima

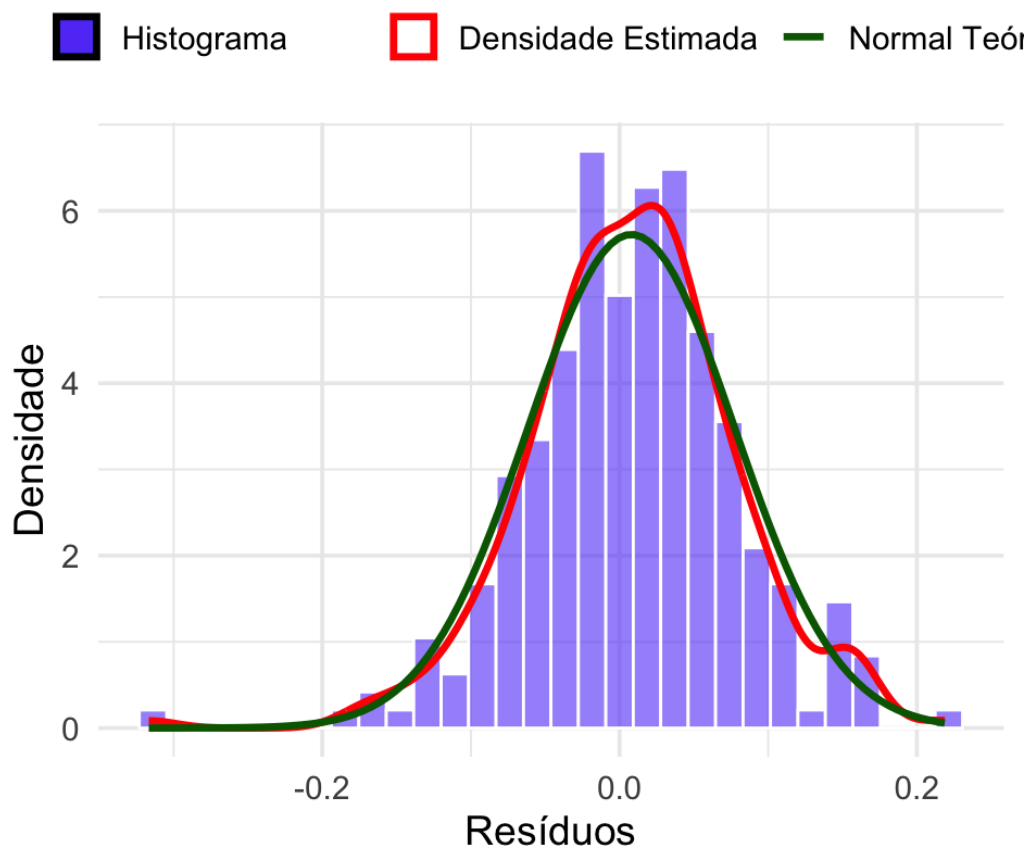


Figura 10: Histograma dos resíduos do modelo SARIMA com densidade estimada (vermelho) e normal teórica (verde).

O QQ-plot (Figura 11) confirma a aproximação à normalidade, com alguns pontos nas caudas escapando das bandas de confiança.



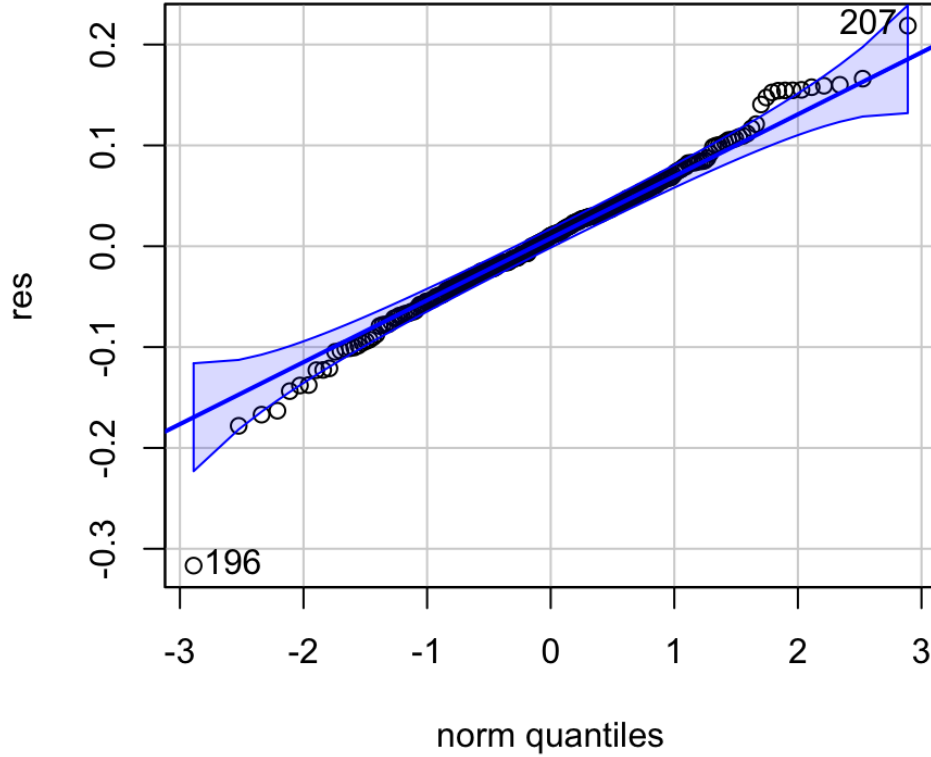


Figura 11: QQ-plot dos resíduos do modelo SARIMA, mostrando leve desvio nas caudas.

**Testes de Normalidade** A Tabela 5 apresenta os resultados dos testes formais de normalidade. Shapiro-Wilk e Jarque-Bera rejeitam normalidade ao nível de 5%, indicando algumas observações atípicas nas caudas, enquanto Lilliefors não rejeita a hipótese de normalidade.

Tabela 5: Resultados dos testes de normalidade nos resíduos

Teste	Estatística	p-valor
Shapiro-Wilk	0.9836	0.0046
Jarque-Bera	30.921	< 0.0001
Lilliefors	0.0340	0.6586

**Raízes dos Polinômios AR/MA/SAR** A análise das raízes confirma a estacionariedade e invertibilidade do modelo, com todos os módulos maiores que 1:

- AR:  $\text{Mod}(\text{roots\_ar}) = 1.3449$
- MA:  $\text{Mod}(\text{roots\_ma}) = 1.0986$
- SAR:  $\text{Mod}(\text{roots\_sar}) \in [1.0253, 1.1167]$

**Conclusão** Os resíduos apresentam comportamento compatível com ruído branco. Apesar de pequenas violações de normalidade nas caudas, o modelo ARIMA(1,1,1)(2,0,0)[12] é adequado para previsão. As raízes confirmam que o modelo é estacionário e invertível, garantindo estabilidade e confiabilidade nas previsões.

#### 4.2.2 Holt-Winters Aditivo

O método de Holt-Winters aditivo foi ajustado à série temporal considerando sazonalidade mensal (período  $s = 12$ ). Os parâmetros de suavização, estados iniciais, desvio padrão dos resíduos e critérios de informação estimados pela função `hw()` foram:

Tabela 6: Modelo Holt-Winters Aditivo

Parâmetro	Valor
Parâmetros de Suavização	
$\alpha, \beta, \gamma$	0.8194, 0.0105, 0.0572
Estados Iniciais	
Nível L	2.6799
Tendência B	0.0149
Sazonal S1	0.0512
Sazonal S2	0.0636
Sazonal S3	0.0595
Sazonal S4	0.0083
Sazonal S5	-0.0444
Sazonal S6	-0.1207
Sazonal S7	-0.1281
Sazonal S8	-0.0650
Sazonal S9	0.0265
Sazonal S10	0.0830
Sazonal S11	0.0246
Sazonal S12	0.0417
$\sigma = 0.0709$	
AIC = 85.86666	AICc = 88.40608    BIC = 146.33274

O modelo Holt-Winters aditivo fornece uma modelagem clara da estrutura de tendência e sazonalidade, apresentando ajustes coerentes com o comportamento histórico da série. Entretanto, devido aos valores muito baixos de  $\beta$  e  $\gamma$ , o modelo tende a suavizar fortemente variações recentes, o que pode limitar sua capacidade de capturar mudanças bruscas ou aumentos recentes na volatilidade. Dessa forma, sua utilidade é maior para previsões suavizadas de médio prazo, enquanto a modelagem de componentes dinâmicos mais complexos é melhor tratada por modelos da classe ARIMA.

As métricas referentes aos resíduos obtidos foram:

Tabela 7: Métricas do Holt-Winters Aditivo

Modelo	RMSE	MAE	MAPE (%)
Holt-Winters Aditivo	0.06868	0.05135	1.3606

### 4.3 Divisão da Série e Avaliação de Desempenho

A série de 21 anos e 7 meses foi dividida da seguinte forma:

- **Treino:** primeiros 20 anos (240 meses)
- **Teste:** últimos 1 ano e 7 meses (19 meses)

As previsões foram geradas a partir dos modelos ajustados previamente e comparadas com os valores observados da base de teste.

Tabela 8: Métricas de Desempenho na Base de Teste

Modelo	RMSE	MAE	MAPE (%)
SARIMA(1,1,1)(2,0,0)[12]	0.07120	0.05800	1.2351
Holt-Winters Aditivo	0.08525	0.06754	1.4574

#### Análise:

- SARIMA apresentou melhor ajuste à estrutura de dependência de curto prazo e tendência.
- Holt-Winters multiplicativo capturou bem a sazonalidade, mas apresentou maior erro geral na previsão da base de teste.

#### Visualização das Previsões:

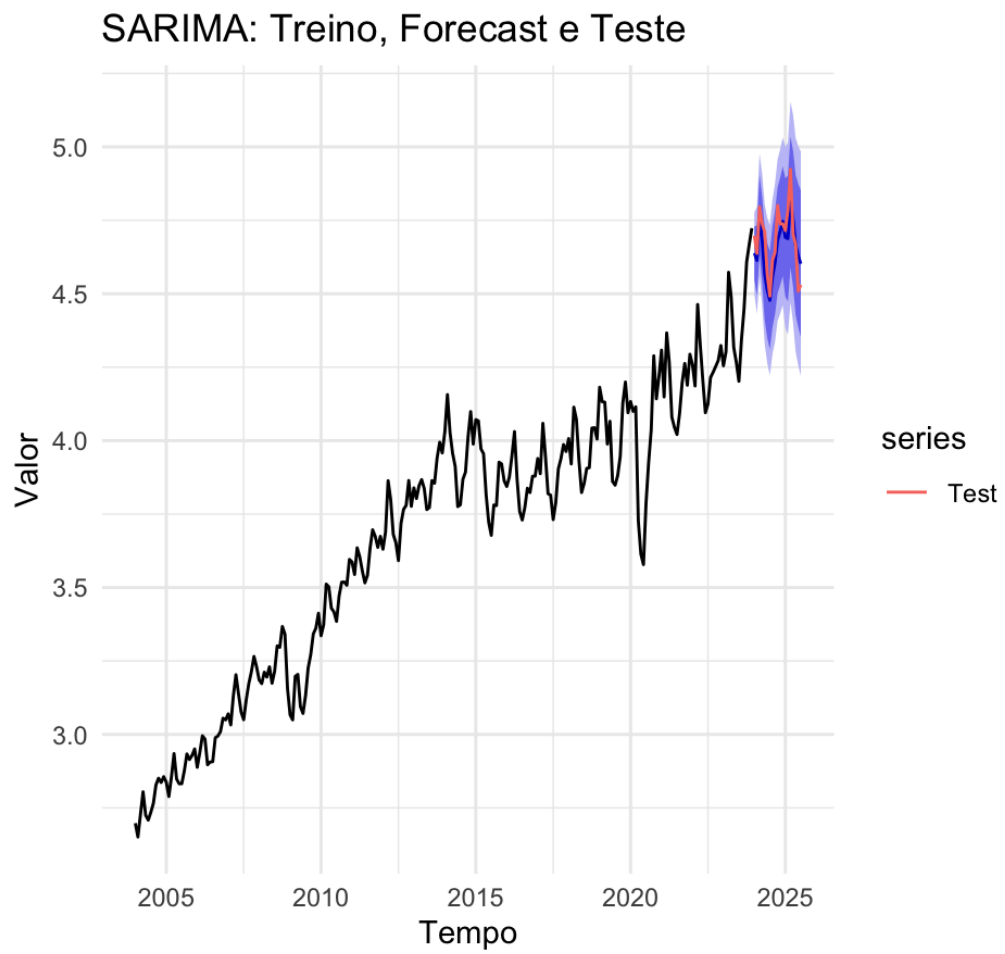


Figura 12: Série histórica e previsão do modelo SARIMA em comparação com a base de teste.

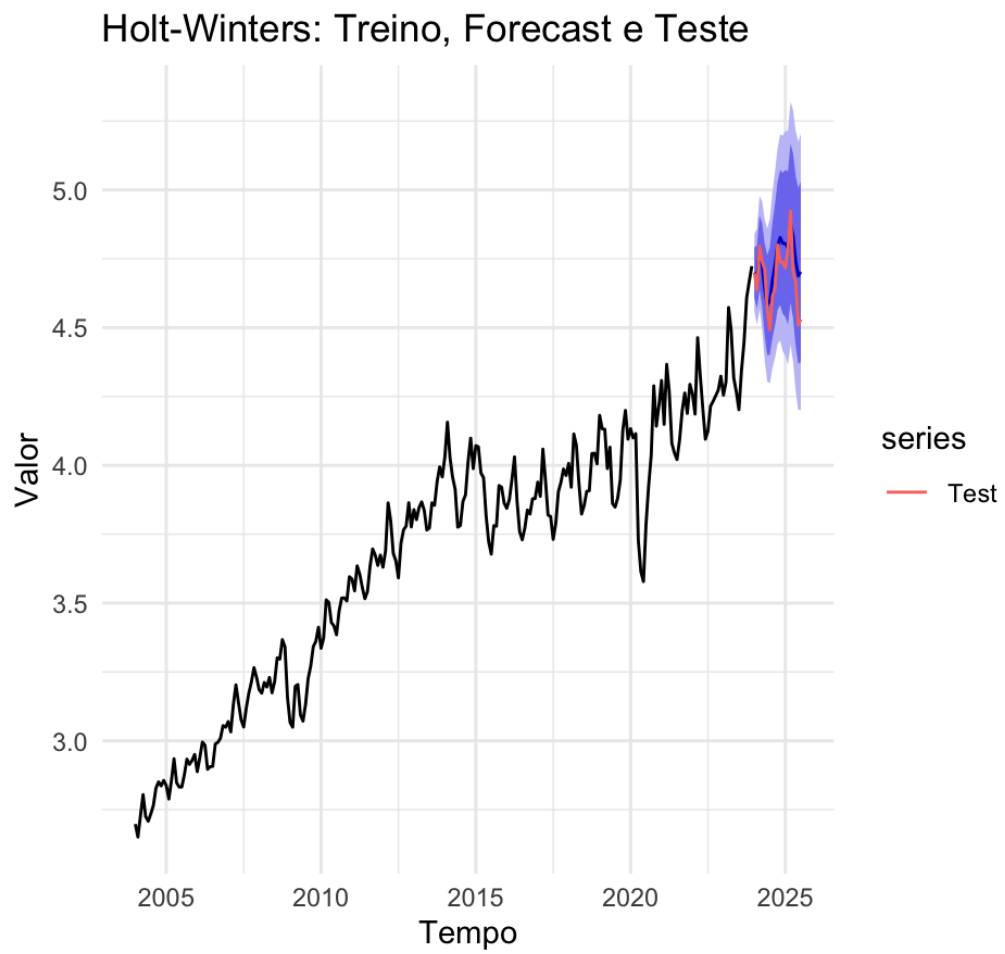


Figura 13: Série histórica e previsão do modelo Holt-Winters aditivo em comparação com a base de teste.

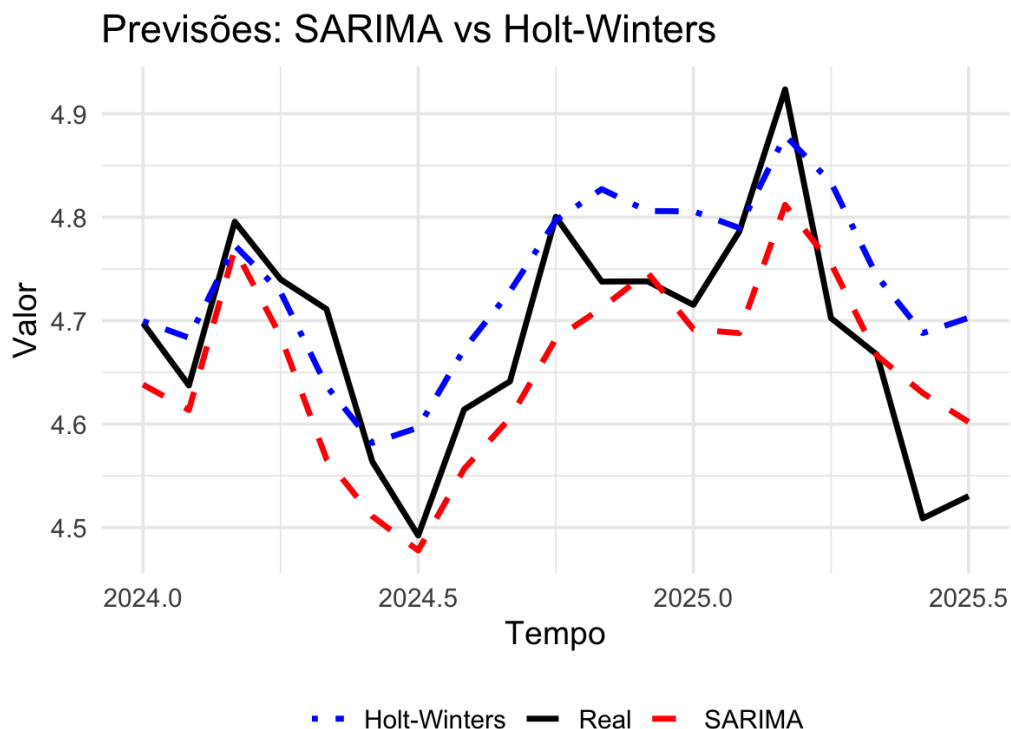


Figura 14: Comparação das previsões do SARIMA e Holt-Winters aditivo com os valores observados da base de teste.

## 5 Conclusão

Este trabalho comparou duas abordagens clássicas de séries temporais — Holt-Winters e SARIMA — aplicadas ao consumo residencial de energia elétrica no Brasil.

O estudo permitiu compreender melhor o comportamento da série, que apresenta tendência de crescimento e sazonalidade anual. Surpreendentemente, o modelo Holt-Winters apresentou desempenho melhor do que o inicialmente esperado, mostrando-se bastante adequado para capturar a estrutura da série.

De forma geral, o trabalho reforça a importância da modelagem de séries temporais para o entendimento de padrões de consumo de energia e demonstra que métodos relativamente simples podem ser eficazes em contextos práticos.